

贝叶斯公式“先验概率”和“后验概率”的应用举例

王世杰, 方晓峰, 赵俊飞

火箭军工程大学基础部数学教研室, 陕西 西安

收稿日期: 2024年10月1日; 录用日期: 2024年10月25日; 发布日期: 2024年11月4日

摘要

在概率论与数理统计中, 贝叶斯公式是一个非常重要的工具, 它用于计算关于复杂事件的条件概率。该研究主要分析了贝叶斯公式在军事案件、生活中以及教学质量评估中的应用, 并创新性地融入了丰富的课程思政元素。让抽象的概率概念变得生动可感, 使学生更好地理解贝叶斯的先验概率和后验概率, 极大地激发了学习者的兴趣与好奇心。

关键词

贝叶斯公式, 条件概率, 先验概率, 后验概率

Application Examples of Bayesian Formulas “Prior Probability” and “Posterior Probability”

Shijie Wang, Xiaofeng Fang, Junfei Zhao

Department of Mathematics, Foundation Department, Rocket Force University of Engineering, Xi'an Shaanxi

Received: Oct. 1st, 2024; accepted: Oct. 25th, 2024; published: Nov. 4th, 2024

Abstract

Bayesian formula is a very important tool in probability theory and mathematical statistics, which is used to calculate the conditional probability of complex events. This study mainly analyzes the application of Bayesian formula in military cases, life and teaching quality evaluation. Let the abstract concept of probability become vivid and sensible, make students better understand Bayesian prior probability and posterior probability, and greatly stimulate learners' interest and curiosity.

文章引用: 王世杰, 方晓峰, 赵俊飞. 贝叶斯公式“先验概率”和“后验概率”的应用举例[J]. 应用数学进展, 2024, 13(11): 4748-4752. DOI: 10.12677/aam.2024.1311456

Keywords

Bayesian Formula, Conditional Probability, Prior Probability, Posteriori Probability

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

贝叶斯公式, 作为概率论中的重要理论, 其诞生源自对条件概率、乘法公式以及全概率公式的深刻洞察与精妙融合。这一公式不仅在理论上构建了先验知识与新信息之间桥梁的坚实框架, 更在实际应用中展现出了无与伦比的灵活性与实用性。全概率公式与贝叶斯公式相辅相成, 共同构成了概率论中不可或缺的重要基石, 而贝叶斯公式尤为突出地展示了如何利用已有的先验信念(先验概率), 在接收新证据后更新这一信念(后验概率)的过程。该研究深入探讨了贝叶斯公式在军事案件、生活中以及教学质量评估中的应用实例, 旨在通过生动的案例分析, 激发学生的学习兴趣。

2. 贝叶斯公式

定理 1 [1] 设实验 E 的样本空间为 S 。 A 为 E 的事件, $B_1, B_2, \dots, B_n$ 为 S 的一个划分, 且 $P(A) > 0$, $P(B_i) > 0$ ($i = 1, 2, \dots, n$), 则

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}, i = 1, 2, \dots, n \quad (1)$$

(1)式称为贝叶斯(Bayes)公式。

证: 由条件概率的定义得:

$$P(B_i|A) = \frac{P(AB_i)}{P(A)}, i = 1, 2, \dots, n \quad (2)$$

由乘法公式得:

$$P(AB_i) = P(A|B_i)P(B_i) \quad (3)$$

由全概率公式得:

$$P(A) = \sum_{j=1}^n P(A|B_j)P(B_j) \quad (4)$$

所以:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^n P(A|B_j)P(B_j)}, i = 1, 2, \dots, n \quad (5)$$

使用贝叶斯公式的基本思路是首先根据加法公式把复杂事件的概率转化为多个简单事件的概率的和, 再根据乘法公式得出在已知复杂事件作为结果已经发生的情况下, 其中某个原因发生的条件概率, 也就是所谓的“后验概率”, 即是“知果索因”的过程[2]。

在贝叶斯统计和概率论中，先验概率是在没有观测到任何数据之前，对某一事件发生的概率的估计或假设。而后验概率则是在考虑了新的观测数据或证据之后，对同一事件发生的概率的重新评估。这种修正过程是通过贝叶斯定理来实现的。贝叶斯定理提供了一种计算后验概率的方法，它结合了先验概率、新的观测数据以及观测数据在所有可能假设下的总体分布。

3. 贝叶斯公式的应用

3.1. 贝叶斯公式在军事上的应用

例 1：飞机坠落在甲、乙、丙、丁四个区域之一，搜救部门判断其概率分别为 0.3、0.2、0.4、0.1，现打算逐个搜索该区域。若飞机坠落在甲、乙、丙、丁四个区域内，被搜索部门发现的概率分别为 0.8、0.7、0.75、0.9。问：首先应该搜索哪个区域？若搜索此区域后，未发现飞机，则此时飞机落入四个区域的概率又是多少呢？[3]

解：设 $A_1 = \{\text{飞机落入甲区域}\}$ ， $A_2 = \{\text{飞机落入乙区域}\}$ ， $A_3 = \{\text{飞机落入丙区域}\}$ ， $A_4 = \{\text{飞机落入丁区域}\}$

(1) 因为 $P(A_1) < P(A_2) < P(A_3) < P(A_4)$ (先验概率)，所以首先应该搜索丙区域。

(2) 设 $\bar{B} = \{\text{首次搜索在丙区域未发现飞机}\}$ ，

由全概率公式得：

$$P(\bar{B}) = \sum_{i=1}^4 P(A_i)P(\bar{B}|A_i) = 0.3 \times 1 + 0.2 \times 1 + 0.4 \times 0.25 + 0.1 \times 1 = 0.7$$

由于事件 $\bar{B}$ 已经发生，因而飞机落入四个区域的概率为

$$\begin{aligned} P(A_1|\bar{B}) &= \frac{P(A_1\bar{B})}{\sum_{i=1}^4 P(A_i)P(\bar{B}|A_i)} = \frac{P(A_1)P(\bar{B}|A_1)}{\sum_{i=1}^4 P(A_i)P(\bar{B}|A_i)} = \frac{0.3 \times 1}{0.7} = \frac{3}{7} \\ P(A_2|\bar{B}) &= \frac{P(A_2\bar{B})}{\sum_{i=1}^4 P(A_i)P(\bar{B}|A_i)} = \frac{P(A_2)P(\bar{B}|A_2)}{\sum_{i=1}^4 P(A_i)P(\bar{B}|A_i)} = \frac{0.2 \times 1}{0.7} = \frac{2}{7} \\ P(A_3|\bar{B}) &= \frac{P(A_3\bar{B})}{\sum_{i=1}^4 P(A_i)P(\bar{B}|A_i)} = \frac{P(A_3)P(\bar{B}|A_3)}{\sum_{i=1}^4 P(A_i)P(\bar{B}|A_i)} = \frac{0.4 \times 0.25}{0.7} = \frac{1}{7} \\ P(A_4|\bar{B}) &= \frac{P(A_4\bar{B})}{\sum_{i=1}^4 P(A_i)P(\bar{B}|A_i)} = \frac{P(A_4)P(\bar{B}|A_4)}{\sum_{i=1}^4 P(A_i)P(\bar{B}|A_i)} = \frac{0.1 \times 1}{0.7} = \frac{1}{7} \end{aligned}$$

此时将飞机落入四个区域的概率修正为

$$P(A_1) = \frac{3}{7}, P(A_2) = \frac{2}{7}, P(A_3) = \frac{1}{7}, P(A_4) = \frac{1}{7} \quad (\text{后验概率})$$

由先验信息可知飞机坠落在甲、乙、丙、丁四个区域的概率分别为 0.3、0.2、0.4、0.1。因为后面增加了新的证据“搜索甲区域还没有发现飞机”，我们可以利用贝叶斯公式计算出在搜索甲区域未发现飞机后，飞机落入甲、乙、丙、丁四个区域的修正概率，修正之后的概率为 $\frac{3}{7}$ ， $\frac{2}{7}$ ， $\frac{1}{7}$ ， $\frac{1}{7}$ 。通过这个概率我们可以判断出下次需要查找的区域。这有助于搜救部门更加高效地分配资源和制定策略，以便尽快找到失

踪的飞机。

思政元素：1) 科学决策与理性分析。案例要求搜救部门在有限的信息下作出最优的搜救决策。这需要我们运用概率论的知识，对各个区域进行科学的评估。通过计算每个区域被搜救部门发现飞机的条件概率，我们可以确定哪个区域是最有可能发现飞机的。这个过程体现了科学决策的重要性，即基于数据和事实进行分析，而不是凭感觉或直觉做决定。2) 透过现象看本质。在确定了首先搜索的区域后，如果未能发现飞机，案例进一步要求我们更新对飞机坠落位置的概率判断。这需要我们深入理解概率的更新机制，即贝叶斯定理的应用。通过这一过程，我们认识到，即使初步的判断未能找到目标，我们也不能轻易放弃，而是要根据新的情况，重新评估各个选项的可能性。这体现了透过现象看本质的能力，即不被初步的结果所迷惑，而是深入分析背后的原因和规律。

3.2. 贝叶斯公式在生活中的应用

例 2：一户人家，男主人外出，突然接到家里的电话说是家里的爱犬金毛找不到了，男主人在安慰了家中的妻子和孩子之后，冷静地根据历史大数据做了如下的基本假设：金毛有 90% 的概率在后院，7% 的概率跑出后院到了森林，3% 的概率被人偷走。男主人告诉家人在家里和后院仔细寻找，家人在后院找了一圈，没有找到金毛，再次焦急地打电话告诉男主人，问下一步如何寻找？假设在后院找过金毛但没找到的概率是 10%。假设事件 A 代表金毛还在后院，事件 B 代表在后院找了金毛，但没有找到。

解：设事件 A_1 、 A_2 、 A_3 分别表示金毛还在后院、金毛跑出后院到了森林和金毛被人偷走。

易知 $P(A_1) = 0.9, P(A_2) = 0.07, P(A_3) = 0.03, P(B|A_1) = 0.1$ (先验概率)

$$P(A_1|B) = \frac{P(B|A_1)P(A_1)}{\sum_{i=1}^3 P(B|A_i)P(A_i)} = \frac{0.1 \times 0.9}{0.1 \times 0.9 + 0.07 \times 1 + 0.03 \times 1} = \frac{9}{19} \approx 47.36\%$$

$$P(A_2|B) = \frac{P(B|A_2)P(A_2)}{\sum_{i=1}^3 P(B|A_i)P(A_i)} = \frac{0.07 \times 1}{0.1 \times 0.9 + 0.07 \times 1 + 0.03 \times 1} = \frac{7}{19} \approx 36.84\%$$

$$P(A_3|B) = \frac{P(B|A_3)P(A_3)}{\sum_{i=1}^3 P(B|A_i)P(A_i)} = \frac{0.03 \times 1}{0.1 \times 0.9 + 0.07 \times 1 + 0.03 \times 1} = \frac{3}{19} \approx 15.79\%$$

此时金毛发生三种情况的概率修正为：

$$P(A_1) = \frac{9}{19}, P(A_2) = \frac{7}{19}, P(A_3) = \frac{3}{19} \text{ (后验概率)}$$

由先验信息可知金毛在后院、跑出后院到森林和被偷走的概率分别为 0.9、0.07、0.03。因为后面增加了新的证据“搜索后院还没有找到金毛”，我们可以利用贝叶斯公式计算出在搜索后院未发现金毛后，金毛在后院、跑出后院到森林和被偷走的修正概率，修正之后的概率为 $\frac{9}{19}$ ， $\frac{7}{19}$ ， $\frac{3}{19}$ 。通过这个概率我们就可以判断出金毛丢失哪个原因发生的概率最大。这有助于家人制定更优的查找策略，以便尽快找到金毛。

思政元素：1) 批判思维和创新。引导学生思考，如果男主人只依赖初始的假设而不根据实际情况调整策略，结果会怎样？这培养了学生的批判性思维。鼓励学生提出新的假设和解决方案，培养他们的创新意识和能力。2) 数学与实际相结合。通过这个例子，让学生理解概率论等数学知识在现实生活中的应用，增强他们学习数学的兴趣和动力。强调数学不仅是抽象的符号和公式，更是解决实际问题的重要工具。

3.3. 贝叶斯公式在教学评估中的应用

例 3: 假设学校需要评估一名学生的阅读能力水平, 将其分为“高”、“中”、“低”三个等级。根据以往经验, 学校对该学生所在年级的学生阅读能力有一个大致的先验分布, 即认为 30% 的学生阅读能力高, 50% 的学生阅读能力中等, 20% 的学生阅读能力低。现在, 一名阅读能力水平高的学生参加了一次阅读测试, 并取得了不错的成绩。已知阅读能力高的学生在测试中取得好成绩的概率为 0.8, 阅读能力中等的学生在测试中取得好成绩的概率为 0.5, 阅读能力低的学生在测试中取得好成绩的概率为 0.3, 学校希望根据这次测试成绩, 更新对该学生阅读能力水平的评估。

解: 事件 A_1, A_2, A_3 表示学生的阅读能力水平高、中、低, 事件 B 表示在阅读测试中取得好成绩

易知: $P(A_1)=0.3, P(A_2)=0.5, P(A_3)=0.2, P(B|A_1)=0.8, P(B|A_2)=0.5, P(B|A_3)=0.3$

由全概率公式得:

$$P(B) = \sum_{i=1}^3 P(B|A_i)P(A_i) = 0.3 \times 0.8 + 0.5 \times 0.5 + 0.2 \times 0.3 = 0.55$$

由贝叶斯公式得:

$$P(A_1|B) = \frac{P(B|A_1) \times P(A_1)}{P(B)} = \frac{0.8 \times 0.3}{0.55} \approx 0.44$$

由先验信息可知有 0.3 的概率认为该学生的阅读能力高。因为后面增加了新的证据“该生参加一次阅读测试取得了不错的成绩”, 我们可以利用贝叶斯公式计算出该生取得不错成绩后该生阅读能力高的修正概率, 修正之后的概率为 0.44。通过这个概率我们就可以判断该生有多大的概率阅读能力高。这有助于学校对学生阅读能力作出更精准的评估, 为学生阅读能力的排名给出参考。

思政元素: 1) 诚信。结合学生参加阅读测试并取得成绩的情境, 讨论诚信在考试中的重要性, 以及如何通过考试成绩进行自我认知和评估。引导学生思考, 如果考试成绩不真实, 会对个人和学校的评估产生什么影响, 从而培养他们的诚信意识和自我反思能力。2) 鼓励自我反思与成长。在评估阅读能力水平的过程中, 学生也可以进行自我反思: 自己的阅读能力如何? 是否还有提升空间? 如何根据测试结果制定改进计划? 这种自我反思的过程有助于学生认识自己、了解自己的优势和不足, 从而制定有针对性的学习计划, 促进个人成长和进步。

4. 小结

贝叶斯公式作为条件概率理论的精髓, 它进一步深化了我们对条件概率的理解。在实际应用中, 贝叶斯公式的有效使用高度依赖于对样本空间的精细划分[4]。通过巧妙地划分样本空间, 我们能够更清晰地识别出不同事件之间的逻辑关系, 进而利用贝叶斯公式进行精确的概率计算。该研究主要分析了贝叶斯公式在军事案件、生活中以及教学质量评估中的应用。让抽象的概率概念变得生动可感, 使学生更好地理解贝叶斯的先验概率和后验概率。

参考文献

- [1] 盛骤, 等. 概率论与数理统计[M]. 北京: 高等教育出版社, 2019.
- [2] 赵明睿. 关于贝叶斯公式及其实际应用[J]. 科学技术创新, 2023(18): 67-70.
- [3] 杨萍, 等. 工程数学工程应用案例及分析[M]. 西安: 西安电子科技大学出版社, 2022.
- [4] 方志平. 例说全概率公式与贝叶斯公式的运用[J]. 中学数学研究, 2023(5): 51-53.